Title

Somite-independent regeneration of the axolotl primary body axis

Authors

Wouter Masselink[1,*], Tobias Gerber[2,3], Francisco Falcon[1,4], Tom Deshayes[1,7], Sofia-Christina Papadopoulos[5,6], Marko Pende[5], Vijayishwer Singh Jamwal[5], Yuka Taniguchi-Sugiura[1], Tzi-Yang Lin[1,4], Thomas Kurth[8], Jingkui Wang[1], Detlev Arendt[3], Ji-Feng Fei[9], Barbara Treutlein[2,10], Elly M. Tanaka[1,*], Prayag Murawala[1,5,6,*]

1. Research Institute of Molecular Pathology (IMP), Vienna BioCenter (VBC), 1030 Vienna Austria.
2. Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany.
3. European Molecular Biology Laboratory (EMBL), Developmental Biology Unit, 69117 Heidelberg, Germany.
4. Vienna BioCenter PhD Program, Doctoral School of the University of Vienna and Medical University of Vienna, 1030 Vienna, Austria.
5. MDI Biological laboratory (MDIBL), Bar Harbor, ME 04609, USA.
6. Clinic for Kidney and Hypertension Diseases, Hannover Medical School, 30625 Hannover, Germany.
7. École Normale Supérieure de Lyon, Université Claude Bernard Lyon 1, Université de Lyon, 69342 Lyon, France.
8. Core Facility Electron Microscopy and Histology. Center for Molecular and Cellular Bioengineering (CMCB), Technology Platform, Technische Universität Dresden, 01307, Dresden, Germany.
9. Department of Pathology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, 510080 Guangzhou, China.
10. Department of Biosystems Science and Engineering, ETH Zurich, Basel 4058, Switzerland.

* Correspondence: wouter.masselink@imp.ac.at, elly.tanaka@imp.ac.at, pmurawala@mdibl.org

Abstract

Primary body axis development is a highly conserved process that proceeds through somitogenesis and further subdivision into dermatome, myotome, and sclerotome. Defects in somitic-clock genes such as *Hes7* lead to vertebral segmentation defects in mice and fish. Here we show that in the axolotl, *Hes7* is also necessary for proper embryonic vertebral segmentation but is, surprisingly, dispensable during axolotl tail regeneration. Using genetic barcoding and fate mapping we found that during regeneration, the somitic derivatives (muscle, cartilage, tendon, fibroblasts) arose from tendon-like, *Lfng*$^{+}$ multi-potent stem cells residing at the myotendonal junction that we term "asomitic stem cells". Throughout homeostasis and regeneration these stem cells display a distinct gene regulatory state compared to developmental progenitors with comparable lineage potential. These observations contrast to axolotl limb regeneration that proceeds via fibroblast dedifferentiation and the subsequent recapitulation of a limb developmental program. Taken together our research shows that divergent strategies are deployed between limb and tail regeneration, and that regeneration of complex body parts does not necessarily involve the complete redeployment of developmental programs.

Main text

The axolotl tail is composed of spinal cord surrounded by segmented vertebrae and muscle (called myomeres) and its regeneration represents a rare example of primary body axis regeneration among vertebrates. While the cellular and molecular basis of spinal cord regeneration has been characterized[1,2] very little is known about how somitic derivatives such as vertebrae are regenerated. During embryonic development, somitic progenitors responding to cues from notochord lay out the architecture of the tail and are the source for vertebrae and myomeres in addition to other connective tissue cell types[3,4]. Regeneration occurs in the absence of notochord but rather blastema cells condense into a cartilage rod and subsequently vertebrae emerge[5]. It is unknown whether myogenic and chondrogenic lineages regenerate from distinct stem cell pools. Also unresolved is whether the tail blastema recapitulates a developmental gene regulatory program as occurs during limb regeneration[6].

Vertebrae regeneration is somite-independent

Previous work has shown that the correct number of myomeres are re-established after tail amputation[7]. We expanded upon this work and found that the number of vertebrae is also correctly re-established (Sup. Fig. 1a-c) (uninjured: 20.10 +/- 1.45 vs regenerated 19.20 +/- 2.10). To assess whether molecular oscillators associated with developmental somitogenesis are involved in tail regeneration, we generated a somitic clock mutant through targeted mutagenesis of *Hes7*. The Hairy and enhancer of split (Hes/Her) family of bHLH transcription factors represent a highly conserved oscillatory gene family involved in somitogenesis[8]. In mice *Hes7* mutants display aberrant somite boundary formation, resulting in vertebral segmentation defects[9,10]. We were unable to recover homozygous *Hes7* mutant axolotls that are apparently embryonic lethal. We could, however, analyze F0 crispants that escaped

embryonic lethality. We injected eggs with Cas9/gRNA complexes representing a mixture of four gRNAs targeting all three exons of the *Hes7* gene sequence (Sup. Fig. 1d) and imaged vertebrae using alizarin red signal in live, 4-5 cm long animals. As in mice, these mutant animals displayed irregular vertebral segmentation and fusions, consistent with a role for *Hes7* in somitogenesis during embryonic development (Fig. 1a-b). When the same tails were amputated, regular vertebral segmentation was found in the regenerate, indistinguishable from non-injected siblings (Fig. 1c-e, Sup. Fig. 1e,f). To ensure that this observation was not due to a positive selection of wildtype and heterozygotic cells in the regenerate, we analysed editing efficiencies at each exonic target site in the unamputated tail close to the amputation plane and in the regenerated tails and found no significant difference in editing efficiency (Fig. 1f). Altogether, these results show that, in the axolotl, *Hes7* is not required for vertebral segmentation during regeneration and raise the possibility that tail regeneration could be somite-independent.

Consistent with the functional analysis of *Hes7*, somites were readily identifiable using H&E staining in the developing axolotl embryo but we found no histological evidence of somites in the tail blastema (Sup. Fig. 1g,h). Another hallmark of somite-mediated patterning is resegmentation where the rostral half of the vertebrae is formed by the caudal part of the somite and vice versa, resulting in a consistent and regular half segment off-set between vertebrae and myomeres. Resegmentation is a highly conserved somite-dependent process, is ancestral to all gnathostomes and is present in axolotl[11,12]. Unamputated axolotl tails showed the expected off-set between the myomeres and vertebrae (Fig.1g-i). In contrast, regenerated tails showed irregularly shaped myomeres[7] and the register between myomeres and vertebrae was randomly distributed after tail regeneration consistent with a lack of somites and resegmentation during regeneration (Fig. 1j-m). In summary, we show that while vertebrae

segment during tail regeneration they do so in the absence of physical somites, somite-dependent resegmentation and a *Hes7* dependent somitic clock.

Multi-potent musculo-skeletal progenitors contribute to tail regeneration

To survey the potency of cells that regenerate the tail, we performed clonal barcode lineage tracing using a CellTag-based strategy adapted to a third-generation foamy virus vector system[13–15]. We generated 3 unique libraries each containing a 9 nt barcode and a defined 6 nt tag (V1, V2, V3) (Fig. 2a, Sup. Fig. 2a-c) associated with the 5' end of a GFP-encoding insert. High titer (Sup. Fig. 2d) viral libraries were used to infect each tail 4 days prior to amputation (V1), as well as 4 (V2), and 7 (V3) days post amputation (dpa) (Fig. 2b-c). We recovered V1, V2, and V3 barcodes from single-cell RNA-sequencing (scRNA-seq) data of low sorted GFP^+ from two independent replicates at 21 dpa (Sup. Fig. 2e-h). To identify putative clones we calculated the Jaccard similarity index of barcodes detected in cells with somite-derivative signatures (tenocytes, intermediate fibroblasts, dermis, fin mesenchyme, chondrocyte precursors, chondrocytes, and myogenic cells) (Fig. 2d-e). This analysis identified clones ranging in size from 2-6 cells which suggests an approximately 6-19% recovery rate when compared to an expected average clone size of 32.89 (V1), 16.91 (V2), and 10.27 (V3) cells, respectively, assuming a blastema cell cycle length of 100 hours. From the 48 clones analyzed 39 clones contained diverse combinations of 2-4 cell types (Fig. 2f). Notably, the myogenic cells showed a clonal relationship with each of the other six cell types. Moreover, computational mixing of both replicates and subsequent clone calling resulted in 47/48 clones being correctly annotated indicating a sufficient diversity of barcodes (Sup. Fig. 2i). Although a single clone containing all seven cell types was not recovered, this data is consistent with the presence of a multipotent musculo-skeletal progenitor in the tail (Sup. Fig. 2j).

The myotendonal junction (MTJ) is the source of multipotent musculo-skeletal progenitors

We next set out to identify and characterize the origin and nature of these progenitors and hypothesized that connective tissue may be a source of such cells. When we grafted vertebrae from a constitutively expressing CAGGs:GFP transgenic animal, we observed minimal contribution, eliminating vertebrae as a cell source (Sup. Fig. 3a-d). We then implemented Cre/loxP genetic fate mapping using two connective tissue-specific Cre-ERT2-expressing drivers (*Col1A2* and *Twist3*) in combination with a *CAGGs:loxP-STOP-LoxP-Cherry* reporter. Recombination was induced by 4OH tamoxifen bathing 10 days prior to tail amputation. Prior to regeneration the *Col1A2* driver broadly labeled the tail connective tissue population (fin mesenchyme, dermal fibroblasts, tenocytes, periskeletal and skeletal cells) (Fig. 3a, Sup. Fig. 4a-d). The *Twist3* driver labeled a smaller subset of tail connective tissue including fin mesenchyme, dermal fibroblasts, and skeletal lineage but not tenocytes that reside at the MTJ (Fig. 3b). After tail amputation, the *Twist3* driver line showed limited participation of the labeled cells to the regenerate (Fig. 3d,f). In contrast, the *Col1A2* driver yielded labeled cells in regenerated fin mesenchyme, tenocytes, and vertebrae as well as in muscle (Fig. 3c,e) indicating this transgenic line drives expression in the progenitors for the somitic derivatives. To assess whether muscle labeling is due to a nonspecific labeling of the myogenic lineage such as satellite cells, we performed sectioning and immunostaining with muscle progenitor marker PAX7. In converted, unamputated tails we did not observe colocalization between the Cherry^{+} cells and PAX7^{+} muscle progenitors (n=100 Cherry^{+} cells/tail, 3 biological replicates) (Sup. Fig. 5a). However, regenerated tails showed the presence of Cherry^{+}/PAX7^{-}, Cherry^{-}/PAX7^{+}, and Cherry^{+}/PAX7^{+} population in addition to Cherry^{+} muscle fibers, suggesting that *Col1A2*-labelled Cherry^{+} cells are hierarchically upstream of at least some PAX7^{+} muscle progenitors (Sup. Fig. 5b).

To molecularly pinpoint the potential source cells, we generated sc-RNAseq data of flow sorted cells from both the *Col1A2* and *Twist3* driver lines (Sup. Fig. 6a-b) and performed a comparative analysis. We annotated six connective tissue cell types (tenocytes, fibroblasts, chondrocytes, fin mesenchyme, dermis and periskeleton) and corroborated cell identities by HCR *in situ* hybridization (Sup. Fig. 6c-h) to localize fin mesenchyme (*Prrx1*$^{+}$), dermal fibroblasts (*Prrx1*^{low}/*Twist3*$^{+}$), periskeletal cells (*Chrdl1*$^{+}$) and chondrocytes (*Sox9*$^{+}$) and tenocytes (*Scx*$^{+}$) at the MTJ. Strikingly, tenocytes were depleted in the Twist3 dataset implicating this cell population as the source cell for regeneration of somitic mesoderm lineages. Examination of the tenocyte gene expression profile revealed expression of *Lfng and Meox1*, two genes expressed during somitic mesoderm development (Fig. 3p). Immunogold labeling of *Col1a2*-labeled cells identified Cherry^{+} cells within the MTJ that interact with muscle fiber ends through collagen deposition (Sup. Fig. 7a,b). Further, whole body light sheet imaging of Col1A2 animals stained for PCNA (proliferative cells) showed that Cherry^{+}/PCNA^{+} cells reside at the MTJ in mature tails (Sup. Fig. 7c, c'). Altogether, our results suggest that asomitic progenitors express tenocyte and somitic progenitor markers, show proliferation and reside at the MTJ. These results implicate *Lfng*$^{+}$/*Meox1*$^{+}$ tenocytes as a source cell population for somitic lineages during tail regeneration.

To assess the trajectories and molecular features by which musculo-skeletal progenitors differentiate during tail regeneration, we performed scRNA-seq analysis on the *Col1A2* labeled lineages at 1 and 2 weeks post amputation (wpa) (Fig. 3g,h). Consistent with our hypothesis trajectory analysis implicated tenocytes as a musculo-skeletal progenitor (Fig. 3i, j) and later blastema time points showed the emergence of the myogenic lineage in the *Col1A2* driver dataset (Fig. 3g). The transcriptional signature of all cell clusters including tenocytes in the blastema showed similarities with the clusters in the mature tail as revealed by Pearson's correlation analysis. This suggests no unique cell cluster emerges during regeneration and

asomitic progenitors represent a resident progenitor of the mature tail. (Fig. 3k,l). Using HCR-ISH of *Lfng*$^{+}$/*Meox1*$^{+}$ as a marker for the asomitic progenitors, we examined their localization during regeneration (Sup. Fig. 7d-i). At 1 wpa we found accumulation of asomitic progenitors throughout the blastema (Sup. Fig. 7d-f) while at 2 wpa myomeres separated by asomitic progenitors emerged (Sup. Fig. 7g-h).

To test whether *Lfng*-expressing cells in the MTJ indeed represent the source cells for regeneration, we generated a *Lfng*-driven Cre-ERT2-expressing transgenic axolotl line. Conversion of the *Lfng* driver prior to regeneration showed that the Cherry expression was absent in the fin mesenchyme and restricted to the inter-myotome population of the medial axolotl tail (Fig. 3m). Upon tail amputation the *Lfng* descendants contributed extensively to the somitic mesodermal lineages of the regenerating tail (Fig 3n,o). Taken together we conclude that tenocyte-like *Lfng*$^{+}$ progenitors, which we term "asomitic stem cells", contribute to regeneration of somitic lineages in the absence of somites.

Asomitic stem cells represent a distinct non-embryonic identity

Our analysis of the *Hes7* mutants along with the scRNA-seq analysis suggested that regeneration of somitic mesoderm lineages differs from the embryonic program. This would be in contrast to limb regeneration where limb connective tissue cells dedifferentiate to a state that is molecularly comparable to a limb mesenchyme progenitor[6]. To further delve into this comparison, we integrated all mature and regeneration timepoints into a single UMAP (Sup. Fig. 8a-a'') and performed scRNAseq on four tail bud stages (stage 25, 28, 30, 35). Comparison of tail development and regeneration datasets showed that genes commonly associated with vertebrate axis elongation, somitogenesis and resegmentation are expressed during embryonic development. We failed to detect increased expression of these genes during tail regeneration compared to uninjured and regenerating tails (Fig. 4a,b). However we did detect expression of

the asomitic stem cell markers Lfng and Meox1 in the combined mature dataset (Fig. 4c, Sup. Fig. 8a-a''). *Uncx4.1* and *Tbx18*, are highly conserved markers of resegmentation. While their expression is detected in the embryo (Sup. Fig. 8b,c), they were not detectable by scRNA-seq (Sup. Fig. 8b) or HCR during tail regeneration. Finally, to assess whether tail blastema cells acquire tail bud-like identities, we performed quadratic programing on previously published limb[16] and the current tail development and regeneration datasets. Our results show that unlike limb blastema cells, tail blastema cells do not acquire an embryonic-like identity (Fig. 4d). In summary, asomitic stem cells have a distinct transcriptional profile and regenerate tail through a transcriptionally distinct route compared to embryonic somite cells.

Conclusions

Here we have shown that regeneration of somitic mesodermal lineages in the tail occurs through processes that are morphologically and molecularly distinct from embryonic development. *Hes7*, a gene associated with the somitic clock, has no apparent role in vertebral segmentation during regeneration. We have identified the source cell of somitic mesodermal lineages during tail regeneration as a *Lfng*$^{+}$/*Meox1*$^{+}$ tenocyte residing at the MTJ, which we term "asomitic stem cell". This appears to be a stem cell that gives rise to multiple mesodermal derivatives during regeneration. Based on this finding, regeneration of somitic mesoderm-derived lineages resembles neural regeneration, where neural stem cells undergo neurogenesis[17], rather than limb regeneration, where fibroblasts dedifferentiate to an embryonic limb bud state.

Previously, tracing of connective tissue in growing larval zebrafish and after muscle injury identified a *Col1A2*-expressing connective tissue cell state with the potential to contribute to muscle differentiation[18,19] and potentially through the activation of MEOX1[20]. As zebrafish does not regenerate its primary body axis, the significance of this progenitor to tail

regeneration and vertebrae segmentation remains unclear. However tail regeneration is a widespread phenomenon across Euteleostomes, in both Sargopterygii and Actinopterygii. Notable examples include the banded knifefish, the African lungfish, the extinct early tetrapod *Microbrachis pelikani* and lizards are all capable of tail regeneration[21–25]. Moreover vertebrae segmentation during regeneration has been documented in the African lungfish, and in *Microbrachis pelikani*, but not to occur in lizards[21,22,24,25]. This suggests that tail regeneration is an ancestral phenomenon, and we speculate that the presence of asomitic stem cells could correlate with the ability to segment vertebrae during primary body axis regeneration.

Conceptually, this work shows that there are two distinctive pathways to forming segmented vertebrae: an embryonic pathway and a post-embryonic regenerative pathway, which has striking implications for regenerative medicine. It will be important to gain more information on the molecular control of this alternative mode of vertebral organogenesis with an outlook to reconstitute this process from adult stem cells in different animal species including mammals.

References

1. Rodrigo Albors, A. *et al.* Planar cell polarity-mediated induction of neural stem cell expansion during axolotl spinal cord regeneration. *eLife* **4**, e10230 (2015).
2. Cura Costa, E., Otsuki, L., Rodrigo Albors, A., Tanaka, E. M. & Chara, O. Spatiotemporal control of cell cycle acceleration during axolotl spinal cord regeneration. *eLife* **10**, e55665 (2021).

3. Teillet, M. A. & Ledouarin, N. M. Consequences of Neural-Tube and Notochord Excision on the Development of the Peripheral Nervous-System in the Chick-Embryo. *Developmental Biology* **98**, 192–211 (1983).
4. Brand-Saberi, B. & Christ, B. 1 Evolution and Development of Distinct Cell Lineages Derived from Somites. in *Current Topics in Developmental Biology* vol. 48 1–42 (Elsevier, 1999).
5. Echeverri, K., Clarke, J. D. W. & Tanaka, E. M. In Vivo Imaging Indicates Muscle Fiber Dedifferentiation Is a Major Contributor to the Regenerating Tail Blastema. *Developmental Biology* **236**, 151–164 (2001).
6. Gerber, T. *et al.* Single-cell analysis uncovers convergence of cell identities during axolotl limb regeneration. *Science (New York, N.Y.)* **362**, eaaq0681-13 (2018).
7. Vincent, C. D., Rost, F., Masselink, W., Brusch, L. & Tanaka, E. M. Cellular dynamics underlying regeneration of appropriate segment number during axolotl tail regeneration. *BMC Dev Biol* **15**, 48 (2015).
8. Krol, A. J. *et al.* Evolutionary plasticity of segmentation clock networks. *Development* **138**, 2783–2792 (2011).
9. Bessho, Y. *et al.* Dynamic expression and essential functions of Hes7 in somite segmentation. *Genes Dev.* **15**, 2642–2647 (2001).
10. Fujimuro, T. *et al.* Hes7 3′UTR is required for somite segmentation function. *Scientific Reports* **4**, 6462 (2014).
11. Piekarski, N. & Olsson, L. Resegmentation in the Mexican axolotl, Ambystoma mexicanum. *J. Morphol.* **275**, 141–152 (2014).
12. Criswell, K. E. & Gillis, J. A. Resegmentation is an ancestral feature of the gnathostome vertebral skeleton. *bioRxiv* **103**, 69–22 (2019).

13. Biddy, B. A. *et al.* Single-cell mapping of lineage and identity in direct reprogramming. *Nature* **564**, 219–224 (2018).
14. Kong, W. *et al.* CellTagging: combinatorial indexing to simultaneously map lineage and identity at single-cell resolution. *Nat Protoc* 1–25 (2020) doi:10.1038/s41596-019-0247-2.
15. Trobridge, G. D. Foamy virus vectors for gene transfer. *Expert Opin Biol Ther* **9**, 1427–36 (2009).
16. Lin, T.-Y. *et al.* Fibroblast dedifferentiation as a determinant of successful regeneration. *Developmental Cell* **56**, 1541-1551.e6 (2021).
17. Lust, K. *et al.* Single-cell analyses of axolotl telencephalon organization, neurogenesis, and regeneration. *Science* **377**, eabp9262 (2022).
18. Ma, R. C., Jacobs, C. T., Sharma, P., Kocha, K. M. & Huang, P. Stereotypic generation of axial tenocytes from bipartite sclerotome domains in zebrafish. *PLoS Genet* **14**, e1007775 (2018).
19. Sharma, P., Ruel, T. D., Kocha, K. M., Liao, S. & Huang, P. Single cell dynamics of embryonic muscle progenitor cells in zebrafish. *Development* dev.178400 (2019) doi:10.1242/dev.178400.
20. Nguyen, P. D. *et al.* Muscle Stem Cells Undergo Extensive Clonal Drift during Tissue Growth via Meox1-Mediated Induction of G2 Cell-Cycle Arrest. *Cell Stem Cell* **21**, 107-119.e6 (2017).
21. Verissimo, K. M. *et al.* Salamander-like tail regeneration in the West African lungfish. *Proceedings of the Royal Society B: Biological Sciences* **287**, 20192939 (2020).
22. Fröbisch, N. B., Bickelmann, C. & Witzmann, F. Early evolution of limb regeneration in tetrapods: evidence from a 300-million-year-old amphibian. *Proc Biol Sci* **281**, 20141550 (2014).

23. Ellis, M. M. *The Gymnotid Eels of Tropical America*. (Published by the authority of the Board of Trustees of the Carnegie Institute, Pittsburgh :, 1913). doi:10.5962/bhl.title.43182.
24. Vonk, A. C. *et al.* Single-cell analysis of lizard blastema fibroblasts reveals phagocyte-dependent activation of Hedgehog-responsive chondrogenesis. *Nat Commun* **14**, 4489 (2023).
25. Fisher, R. E. *et al.* A Histological Comparison of the Original and Regenerated Tail in the Green Anole, *Anolis carolinensis*. *The Anatomical Record* **295**, 1609–1619 (2012).
26. Woodcock, M. R. *et al.* Identification of Mutant Genes and Introgressed Tiger Salamander DNA in the Laboratory Axolotl, Ambystoma mexicanum. *Sci Rep* **7**, 6 (2017).
27. Khattak, S. *et al.* Germline transgenic methods for tracking cells and testing gene function during regeneration in the axolotl. *Stem Cell Reports* **1**, 90–103 (2013).
28. Kawaguchi, A. *et al.* Chromatin states at homeoprotein loci distinguish axolotl limb segments prior to regeneration. *bioRxiv* 2022.11.14.516253 (2022) doi:10.1101/2022.11.14.516253.
29. Aulehla, A. *et al.* A β-catenin gradient links the clock and wavefront systems in mouse embryo segmentation. *Nat Cell Biol* **10**, 186–193 (2008).
30. Fei, J.-F. *et al.* Efficient gene knockin in axolotl and its use to test the role of satellite cells in limb regeneration. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 12501–12506 (2017).
31. Fei, J. F. *et al.* Application and optimization of CRISPR-Cas9-mediated genome engineering in axolotl (Ambystoma mexicanum). *Nat Protoc* **13**, 2908–2943 (2018).
32. Kroll, F. *et al.* A simple and effective F0 knockout method for rapid screening of behaviour and other complex phenotypes. *eLife* **10**, e59683 (2021).

33. Riquelme-Guzmán, C. & Sandoval-Guzmán, T. Methods for Studying Appendicular Skeletal Biology in Axolotls. in *Salamanders* (eds. Seifert, A. W. & Currie, J. D.) vol. 2562 155–163 (Springer US, New York, NY, 2023).
34. Pende, M. *et al.* A versatile depigmentation, clearing, and labeling method for exploring nervous system diversity. *Sci Adv* **6**, (2020).
35. Kurth, T. Immunocytochemistry of the amphibian embryo--from overview to ultrastructure. *Int J Dev Biol* **47**, 373–383 (2003).
36. Kurth, T. *et al.* Electron Microscopy of the Amphibian Model Systems Xenopus laevis and Ambystoma mexicanum. in *Methods in Cell Biology* vol. 96 395–423 (Elsevier, 2010).
37. Wagner, F. *et al. Human Photoreceptor Cell Transplants Integrate into Human Retina Organoids*. http://biorxiv.org/lookup/doi/10.1101/2022.08.09.500037 (2022) doi:10.1101/2022.08.09.500037.
38. Choi, H. M. T. *et al.* Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* **145**, dev165753 (2018).
39. Lovely, A. M., Duerr, T. J., Stein, D. F., Mun, E. T. & Monaghan, J. R. Hybridization Chain Reaction Fluorescence In Situ Hybridization (HCR-FISH) in Ambystoma mexicanum Tissue. in *Salamanders* (eds. Seifert, A. W. & Currie, J. D.) vol. 2562 109–122 (Springer US, New York, NY, 2023).
40. Bensimon-Brito, A. *et al.* Revisiting in vivo staining with alizarin red S - a valuable approach to analyse zebrafish skeletal mineralization during development and regeneration. *BMC Dev Biol* **16**, 2 (2016).
41. Replogle, J. M. *et al.* Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nat Biotechnol* **38**, 954–961 (2020).

42. Khattak, S. *et al.* Foamy virus for efficient gene transfer in regeneration studies. *BMC Dev Biol* **13**, 9 (2013).

43. Wallace, H. & Maden, M. The cell cycle during amphibian limb regeneration. *Journal of Cell Science* **20**, 539–547 (1976).

44. Schreckenberg, G. M. & Jacobson, A. G. Normal stages of development of the axolotl, Ambystoma mexicanum. *Developmental Biology* **42**, 391–399 (1975).

45. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* **9**, 171–181 (2014).

46. Nowoshilow, S. *et al.* The axolotl genome and the evolution of key tissue formation regulators. *Nature* 1–21 (2018) doi:10.1038/nature25458.

47. Schloissnig, S. *et al.* The giant axolotl genome uncovers the evolution, scaling, and transcriptional control of complex gene loci. *Proceedings of the National Academy of Sciences* **118**, e2017176118 (2021).

48. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

49. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**, 411–420 (2018).

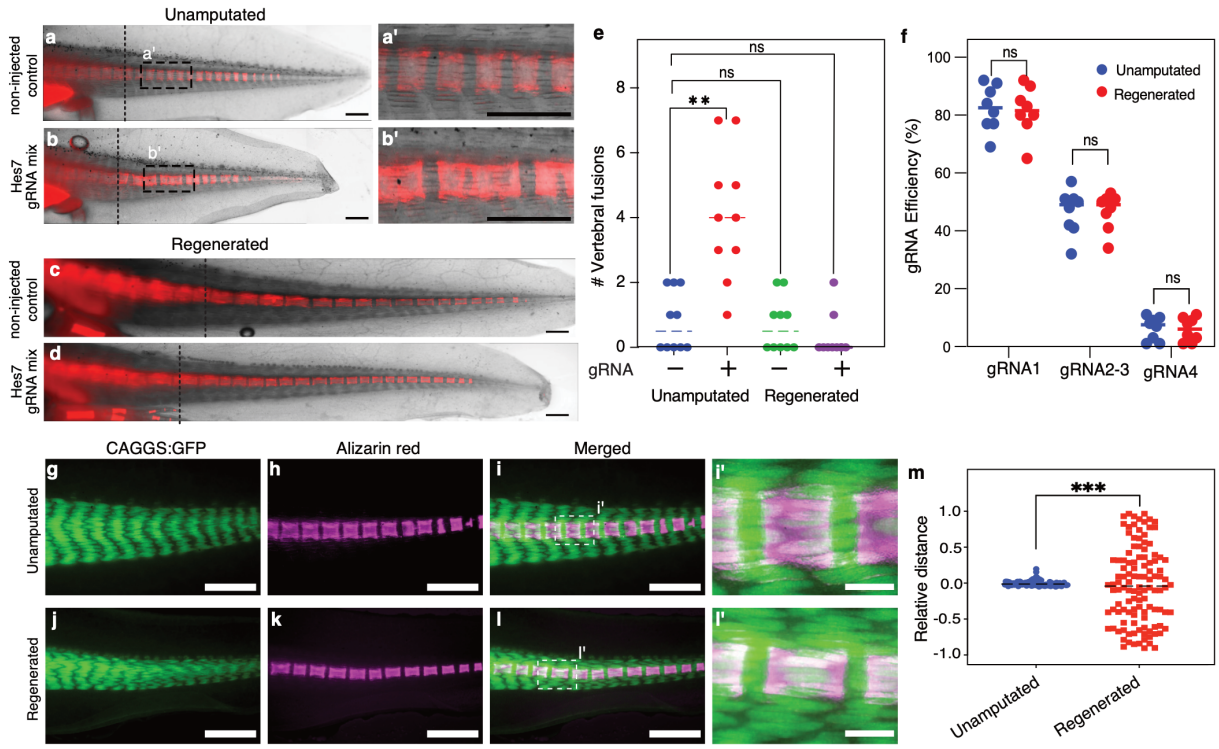
50. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* **16**, 1289–1296 (2019).

51. Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol* **38**, 1408–1414 (2020).

52. Wolf, F. A. *et al.* PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol* **20**, 59 (2019).

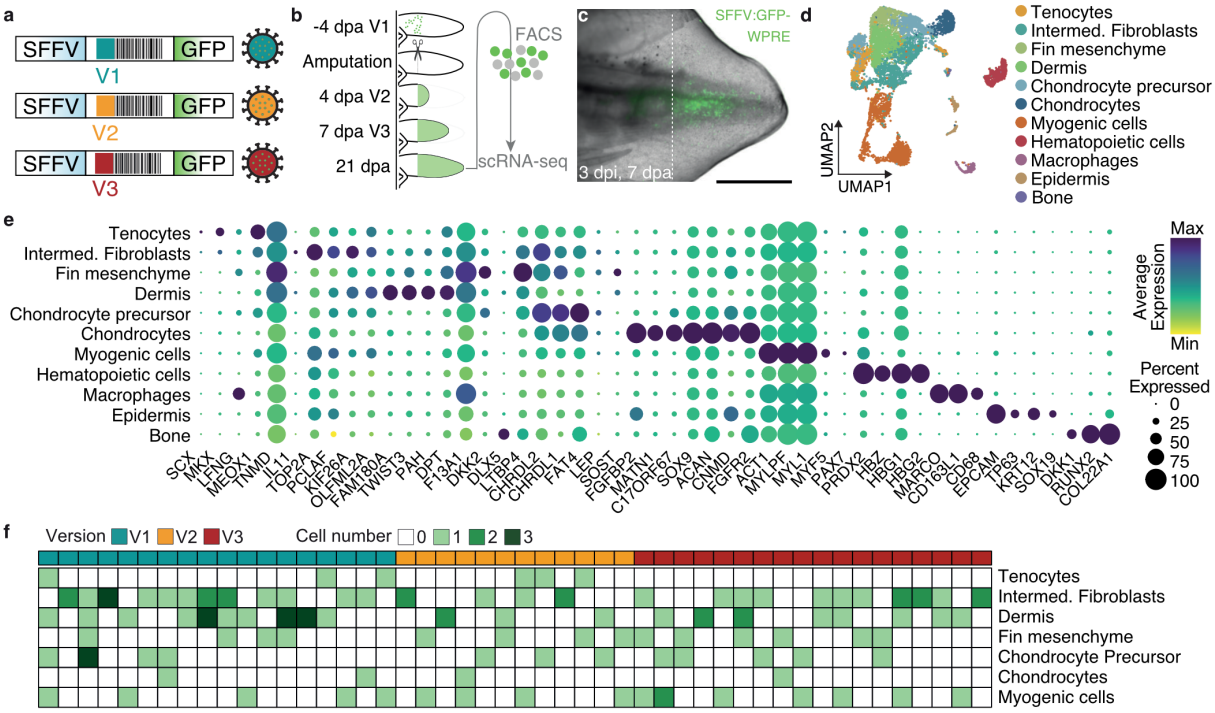
53. Angerer, P. *et al.* destiny: diffusion maps for large-scale single-cell data in R. *Bioinformatics* **32**, 1241–1243 (2016).

Figure 1:



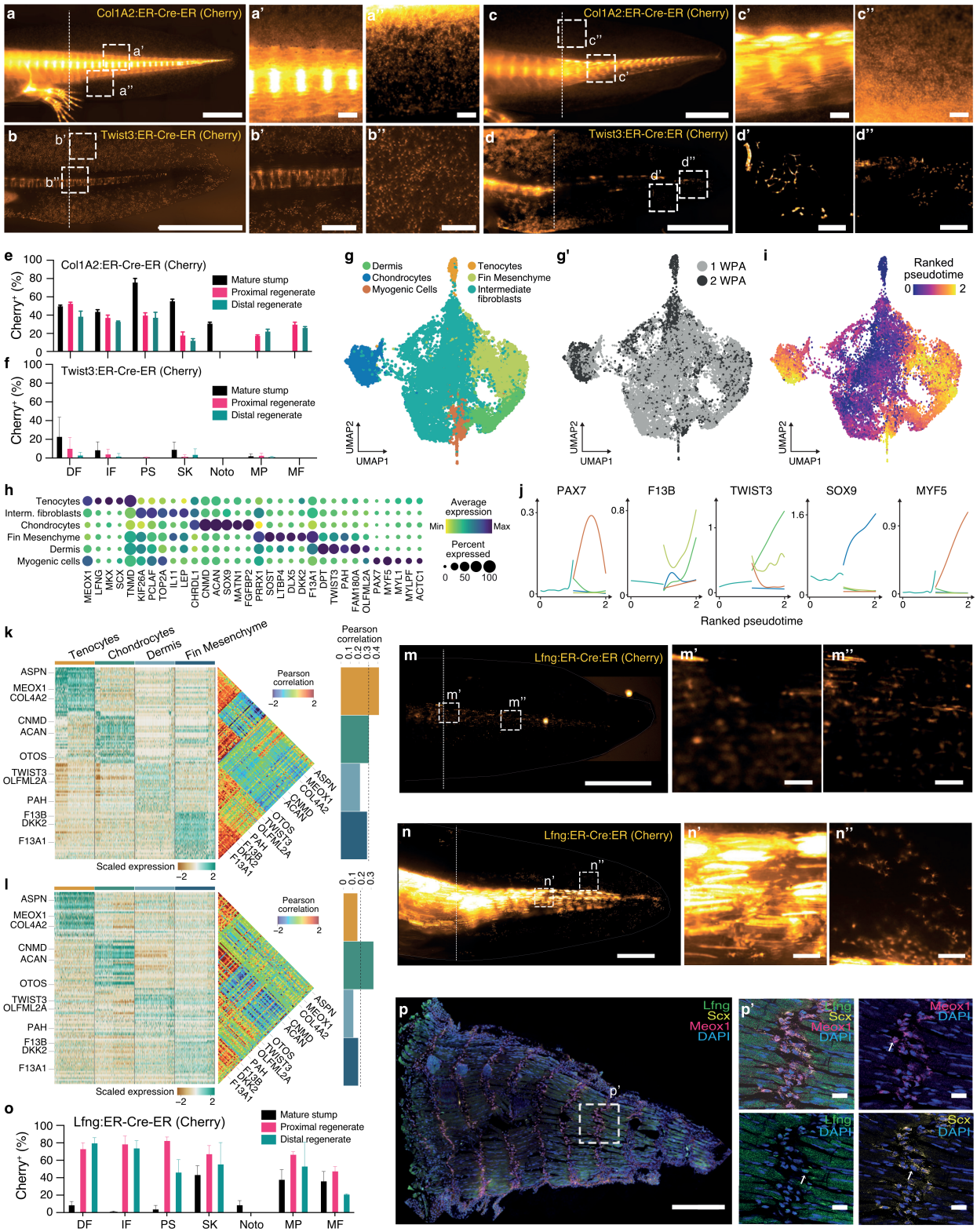
a-d, Stereoscopic images of alizarin red labelling of segmented vertebrae in uninjured non-injected control (a) *Hes7* gRNA mix injected axolotl (b), regenerated non-injected control (c) and regenerated *Hes7* gRNA mix injected axolotl (d). **e**, Scoring of vertebral fusion effects in the tail (a-d) (n=10). **f**, Quantification of gRNA editing efficiency of unamputated and regenerated tail (b and d) (n=8). **g-i**, Stereoscopic images of a *CAGGS:GFP* transgenic axolotl labelled with Alizarin red in uninjured (g-i) and regenerated (j-l) tail. **m**, Quantification of muscle vertebrae off-set in the tail of i and l (n=112). Statistical analysis: e: Kruskal-Wallis test, followed by Dunn's test. f: one-way Anova, followed by Šidák correction. m: F-test. ns: not significant, ** $p < 0.01$, *** $p < 0.001$, Scalebar: a-b': 500 µm. g-l: 1000 µm.i', l': 200 µm.

Figure 2:



a, schematic overview of foamy virus barcode labelling vectors. **b**, Experimental overview of infection, amputation and harvesting timepoints. **c**, Representative image of a 7 dpa foamy virus infected axolotl tail infected at 4 dpa. **d-e**, UMAP and dotplot representation of different clusters identified in GFP^{+} cells from foamy virus infected tails combining two independent replicates. **f**, Matrix based depiction of identified clones with a Jaccard similarity index of ≥0.45 (barcodes present in ≥ 2 cells and ≥ 2 UMIs per cell). Each row represents a cell-type. Each column represents a clone.

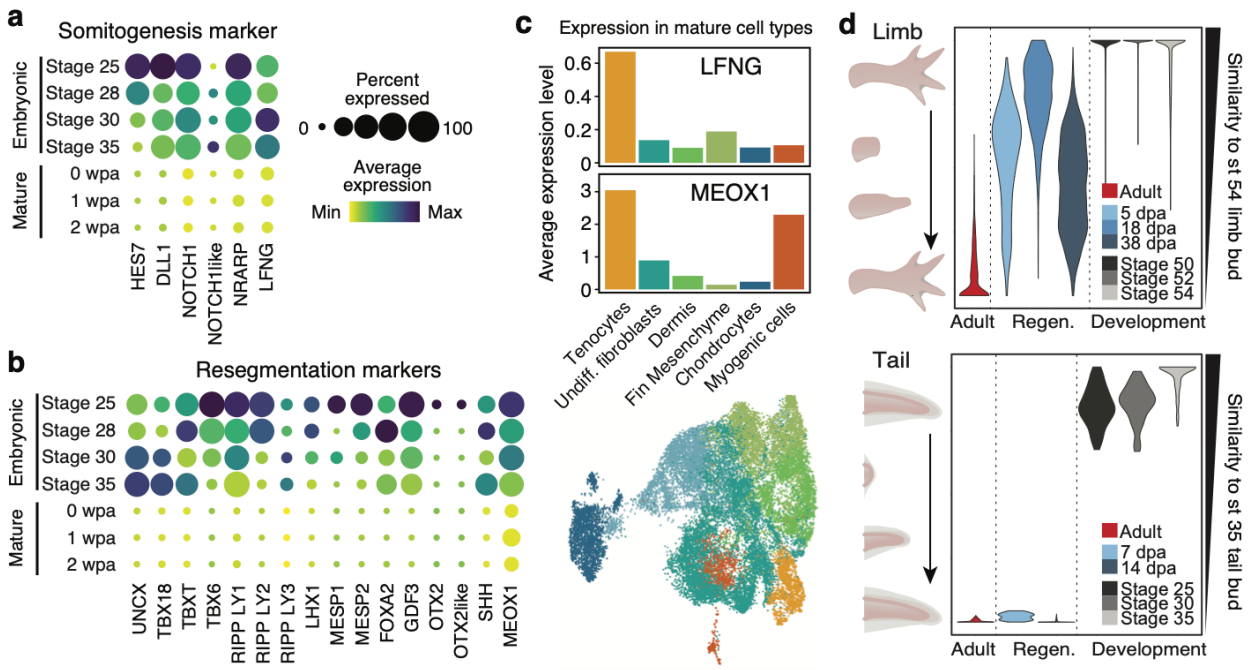
Figure 3:



a-d, Labelling of 4OHT treated *Col1a2:ER-Cre-ER;Caggs:lp-Cherry*, and *Twist3:ER-Cre-ER;Caggs:lp-Cherry* axolotl tail before amputation (a,b), and at 6 wpa (c,d). **e,f**, Quantification of cell type distribution at three different positions (mature stump, proximal and distal

regenerate, Sup. Fig. 3g) of the *Col1a2:ER-Cre-ER;Caggs:lp-Cherry* (e) and *Twist3:ER-Cre-ER;Caggs:lp-Cherry* (f) regenerated tail. **g,g'**, UMAP integration of cells from 1 and 2 wpa regenerated tail colored by cell type identity (left) and by timepoint (right). **h**, Dotplot representation of cell type marker genes (cols) across the cell type (rows) identified in g. **I,** Pseudotime estimates by a diffusion map analysis are visualized on the UMAP embedding. **j,** Line plots represent the smoothed gene expression across cells along their pseudotime separated by cell type, respectively. **k**, Heatmap for the top 30 markers (rows) of each cell type (cols) in the uninjured scRNA-seq data (see Sup. Fig. 6a-b). The genes were used as input for a gene correlation analysis visualized as heatmap with average values within each cell type being summarized as barplot (right). **l**, Genes in (k) are similarly presented for the blastema data showing a similar degree of gene coexpression in the cell types during regeneration. **m-n**, Labelling of 4OHT treated *Lfng:ER-Cre-ER;Caggs:lp-Cherry* axolotl before amputation (m), and at 6 wpa (n). **o**, Quantification of cell type distribution of the regenerated *Lfng:ER-Cre-ER;Caggs:lp-Cherry* tail, including the mature stump, proximal and distal regenerates. DF: dermal fibroblasts, IF: interstitial fibroblast, PS: peri-skeleton, SK skeleton, Noto: notochord, MP: muscle progenitor, MF: muscle fiber. (Sup. Fig. 3g). **p**, HCR-ISH for *Lfng*, *Scx*, and *Meox1* on mid-saggital sections of the mature tail. Scale bars: a-d, p: 500 µm a'-d'', p': 50 µm, m-n: 1000 µm, m'-n'': 100 µm.

Figure 4:



a,b, Dotplot representation comparing the expression of various somitogenesis (a) and resegmentation (b) markers between tail bud, mature tail and tail blastema datasets. **c,** Tissue specific expression of *Lfng* and *Meox1* in a combined dataset of all mature tail samples. UMAP of all mature tail samples (details are presented in Sup. Fig. 8a-a''). **d,** Comparative quadratic programming shows that limb cells acquire embryonic limb mesenchyme identity during limb regeneration (from Lin et al[16]) but tail cells do not acquire tail bud identity.

Materials and Methods (3378 words)

Animal Husbandry, transgenesis, and 4-OH tamoxifen treatment

Axolotls (*Ambystoma mexicanum*) were bred in IMP and MDIBL facilities. All animal handling and surgical procedures were carried out in accordance with local ethics committee guidelines. Animal experiments were performed as approved by the Magistrate of Vienna and IACUC committee of MDIBL. "White" refers to a non- transgenic d/d strain of axolotl that has leucistic skin due to the defect in melanocytes migration during embryonic development and carries *Edn3* mutation[26]. Animals were anaesthetised in 0.03% benzocaine (Sigma) before

amputation and surgery. Axolotl husbandry, transgenesis and 4-hydroxytamoxifen (4OHT) treatment are done as previously described[43].

tgISceI(Mmu.*Col1a2:TFPnls-T2a-ER^{T2}-Cre-ER^{T2}*)^{Etnka}, tgISceI(*CAGGs:LoxP-GFP-LoxP-Cherry*)^{Etnka} and tgISceI(*Caggs:LoxP-STOP-LoxP-Cherry*)^{Etnka} transgenic lines are described previously[6,27,28]. To generate *Lfng* transgenic animals, mouse *Lfng*[29] driver element was cloned at the 5′ end of the GFPnls-T2A-ER^{T2}-Cre-ER^{T2} (ER-Cre-ER) cassette with flanking ISceI sites. *Twist3* transgenic animal was generated by knock-in methodology described previously[30]. *Twist3* is a single exon gene. First, to generate a targeting construct, genomic seq containing 450 bp upstream seq of 5'UTR of *Twist3* along with its 5'UTR and ORF (without stop codon) (total 1230 bp) was amplified from axolotl genomic DNA and was cloned at the 5' end of MCS-T2a-GFPnls-T2a-ER^{T2}-Cre-ER^{T2} base vector to generate 450bp_upstream_seq-5'UTR-Twist3ORF(without stop codon)-T2a-GFPnls-T2a-ER^{T2}-Cre-ER^{T2} construct. Next, embryos were co-injected with the targeting construct, CAS9 protein and gRNA (GCTGCAGCCCATGCAATC-TGG) with a target site 295 bp upstream of the ORF to create dsDNA breaks in both genome and targeting construct. Embryos were screened for nuclear GFP fluorescence at the 3 cm stage and grown up as potential breeders. The successful germ-line transmission was evaluated by nuclear GFP fluorescence in dermal fibroblasts across the body and a line was established. Constructs used for generating transgenic axolotl driver lines are available at Addgene (200104, 200105, 200106 and 200107).

Generation of *Hes7* mutants

Hes7 mutant animals were generated by co-injecting freshly fertilised d/d eggs with RNP complexes containing a cocktail of 4 different guide RNAs targeting all 3 exons of *Hes7* using standard methods[31,32]. gRNA was prepared using primer templates and *in vitro* RNA synthesis. gRNA efficiency was assessed by targeted amplification of gRNA target sites

compatible with Interference of CRISPR Edits (ICE) analysis. In brief gDNA was prepared by incubating 1- 2 mm of tail tissue with 200 µL of 50mM NaOH at 95°C for 15-20 minutes, and subsequently neutralized by mixing with 50 µL of 1M TRIS-in 50 mM NaOH. Samples were centrifuged at 15000 RCF for 6 minutes and supernatant was subsequently used for PCR. 1 µL of gDNA was used to seed 50 µL PCR reactions and cleaned up for sanger sequencing using magnetic beads. Sanger sequencing was performed using a primer that is internal to the amplicon. gRNA efficiency was determined through ICE analysis software (Synthego, Redwood City). Primers are listed in Table 1.

Histology

Embryos were fixed in Carnoy's solution, blastema and mature tissues were fixed in 4% PFA. H&E staining was done using an Epredia Gemini AS Automated Slide Stainer (Fisher Scientific, Waltham) with standard protocols. Stained slides were imaged using a Pannoramic 250 FLASH II digital scanner (3DHistech Ltd., New Jersey).

Lineage tracing and analysis

To perform Cre-LoxP lineage tracing experiments, Cre driver lines were crossed with either tgISceI(*CAGGs:LoxP-GFP-LoxP-Cherry*)^{Etnka} or tgISceI(*CAGGs:LoxP-STOP-LoxP-Cherry*)^{Etnka}. Animals and double transgenic animals were grown up to 4 cm snout to tail size. Animals were treated three times at alternate days with 2 µM 4OHT by bathing to obtain indelible Cherry labelling of cells. Labelled animals were amputated 10 days post-conversion (from the first 4OHT treatment), near the 6th myotome from the cloaca and contributions of *Cherry*$^{+}$ cells to the regenerate were accessed by stereoscopic imaging of a regenerated tail every week for the next 6-weeks using an Axio zoom.V16 (Zeiss, Jena) equipped with an Orca-Flash4.0 Camera (Hamamatsu Photonics, Hamamatsu) and an X-cite Xylis LED illuminator

(Excelitas Technologies, Waltham). At the end of 6 weeks, tails were harvested near cloaca and IHC or HCR analysis was performed.

Antibodies

The primary antibodies used in these studies are rabbit anti-RFP (Rockland# 600-401-379, 1:100), mouse anti-COL1A2 (DSHB # Sp1.D8, 1:100), mouse anti-PCNA (Calbiochem#NA03, 1:500), mouse anti-PAX7 (DSHB#Pax7, 1:100), mouse anti-MHC (monoclonal antibody 4A1025, a kind gift from S. Hughes), goat anti-SOX9 (R&D#AF3075, 1:100), Rabbit anti-SOX9 (Chemicon#AB5535, 1:1000), rabbit anti-PRRX1[6] (1:200) rabbit anti-Laminin (Merck, #L9393 1:200). Secondary antibodies used in these studies were procured from Thermo Fisher Scientific (Waltham).

Immunohistochemistry and microscopy

IHC staining was done as previously described[6]. Briefly 10 µm sections were permeabilized using phosphate-buffered saline (PBS) containing 0.3% Tween-20 and blocked with PBS with 0.3% Triton-X100 and 2% normal hose serum (Vector Laboratories # S-2000. Slides were incubated O/N with primary antibodies in blocking buffer in a humidified chamber at room temperature and subsequently incubated with secondary antibodies and Hoechst at a concentration of 0.5 µg/mL. Mosaic images of sections were acquired on an Axio Imager.Z2 (Zeiss, Jena) equipped with an Orca-Flash4.0 Camera (Hamamatsu Photonics, Hamamatsu) using a 20x/0.8 plan-apochromat objective. Alcian Blue / Alizarin Red staining was performed as previously described[33]. Stained axolotls were photographed using a Nikon D40 DSLR camera with an AF-S DX NIKKOR 18-55mm f/3.5-5.6G VR II objective (Nikon, Tokyo) and an iPad Air (4th gen) (Apple, Cupertino) as a backlight illuminator.

Identification of different cell types based on histology

Collagen 1 alpha 2 (COL1A2) is an extracellular matrix (ECM) protein and in axolotl, it is expressed in dermal fibroblasts and skeleton cells (28). Staining tail sections with COL1A2 antibodies showed broad labelling of the basal lamina (that demarcates dermal fibroblasts and basal keratinocytes), skeletal and periskeletal structures, as well as fin mesenchyme. Thus, based on co-staining of PRRX1 and COL1A2 along with Laminin and Pax7 of sister sections, we identified the following five distinct CT subtypes (Supplementary Fig. 4). 1 "dermal fibroblasts (DF)" - cells underneath the basal lamina that are $COL1A2^{+}$ and PRRX1med. 2. "Interstitial fibroblasts (iF)" cells scattered across the fin mesenchyme that are $COL1A2^{+}$ and $PRRX1^{high}$ or $PRRX1^{med}$. 3."Periskeletal cells (PS)" - cells at the periphery of skeleton that are $COL1A2^{+}$ and $PRRX1^{high}$ or $PRRX1^{med}$, 4. "Skeletal cells (Sk)" - cells which reside inside cartilaginous structure and are $COL1A2^{+}$ and $PRRX1^{low}$. 5. "Notochordal cells (Noto)" – large vacuole containing cells within the embryonic notochord. 6. "Muscle progenitors (MP)" - $PAX7^{+}$ cells scattered in between the muscle fibers, and 7. "Muscle fibers (MF)" – MHC^{+} or $Laminin^{+}$ cells within the myotome region.

Whole mount light sheet imaging of PCNA labelled cells

In brief axolotls samples were cleared and labelled using the deep-clear protocol followed by imaging on a custom built light-sheet system[34]. Samples were fixed in 4 % PFA at 4°C overnight. Specimens were washed several times with PBS at room temperature to remove fixative. After PBS washes, specimens were incubated in Solution-1 at 37°C for 1 day under gentle shaking. After the Solution-1 treatment step of tissue clearing and five short PBS washes at room temperature, samples were treated with 10 % sheep serum at room temperature for several hours. Samples were then incubated with primary antibodies in 5 % sheep serum at 4°C

(gentle shaking) followed by additional wash steps followed by incubation with secondary antibodies in 5 % sheep serum at 4°C (gentle shaking).

Solution-1 consists of 8% THEED (Sigma-Aldrich, 87600-100ML), 5% Triton® X 100 (Roth, 3051.2), 25% Urea (Roth, X999.2) and 5% CHAPS (Hopax) in dH_2O. Solution-2 was prepared by mixing 50% Meglumine diatrizoate (Sigma-Aldrich M5266) in PBS (pH 8.5), and refractive index was adjusted to 1.45.

Images were acquired using a custom build light-sheet system equipped with a Olympus Plan Achromat 1.0/0.25 objective (N1564200). Sample were illuminated with LightHUB 4 system (Omicron, Germany) containing following single mode laser: 120mW 405nm LuxX laser, 150mW 488nm LuxX, 150mW 561nm Cobalt, 100mW 594nm Cobalt, 140mW 647nm LuxX laser. Images were recorded in 16bit quality as Tiff, with the Kinetix camera (Teledyne Photometrics, USA) 3200x3200px; pixel size 6.5 µm. Images were processed with AMIRA (Thermo Scientific, USA) rendering software.

Immunogold labelling and Electron Microscopy

Samples were fixed in freshly prepared 4% PFA. Longitudinal 50 mm thick vibratome sections were prepared on a Leica VT 1200 vibratome and sections with target cells were selected based on their Cherry fluorescence. Vibratome sections were subjected to pre-embedding immunogold labelling using Nanogold and silver enhancement[35–37]. In brief, samples were blocked and permeabilized in 20% normal goat serum (NGS) / PBS / 0.1% Saponin for 2 hours at room temperature followed by incubation in primary antibody (rabbit anti-RFP, Rockland, # 600-401-379, 1:100) in 20% NGS/0.05% Saponin for 2 days at room temperature. After washes in 20% NGS/0.05% Saponin the samples were incubated with goat-anti-rabbit Nanogold (Fab-fragments, Nanoprobes, 1:50) in 20% NGS/0.05% Saponin overnight at room temperature. After final washes in 20% NGS/0.05% Saponin the samples

were postfixed in 1% glutaraldehyde in PBS, washed in water, silver enhanced using the SE-Kit (Aurion, 1 hour incubation time), followed by washes in water, post-fixation in 1% osmium tetraoxide/water (1 hour on ice), washes in water, *en bloc* contrasting with 1% uranyl acetate (1h, on ice), washes in water, and dehydration in a graded series of ethanol (30%, 50%, 70%, 90%, 96% ethanol/water, 3x 100% ethanol on a molecular sieve). The samples were infiltrated in mixtures of the Epon substitute EMbed 812 with Ethanol (1+2, 1+1, 2+1, 2x pure epon), flat embedded on the surface of an empty epon dummy block, and cured overnight at 65°C. The region of interest (ROI) was identified on semithin sections (1 mm) stained with 1% toluidine blue/0.5% borax and the ROI in the block was trimmed for ultrathin sectioning. 70 nm sections were prepared using a diamond knife (Diatome) and the Leica UC6 ultramicrotome (Leica Microsystems, Wetzlar). The sections were mounted on formvar coated slot grids and contrasted with 4% uranyl acetate/water for TEM. Finally, sections were imaged with a Jeol JEM1400Plus transmission electron microscope running at 80kV acceleration voltage and equipped with a Ruby digital camera (Jeol, Tokyo).

HCR-ISH

Samples were fixed in freshly prepared 4% PFA overnight at 4˚C. Prepared for cryo-sectioning and sectioned at 80 µm thickness. HCR staining was performed according to the previously published protocol[38,39] The probes were designed using probe generator software from the Monaghan lab (https://probegenerator.herokuapp.com/ - Not functional anymore) or by the MDIBL computational BioCore team and probes were ordered from IDT (Coralville) as oligo pools. All other HCR reagents were ordered from the Molecular instruments (Los Angeles).

HCR-ISH on slides images were acquired using a point scanning confocal unit (LSM 980, Carl Zeiss Microscopy, Germany) on a Zeiss Axio Examiner Z1 upright microscope stand

(ref: 409000-9752-000, Carl Zeiss Microscopy, Germany) equipped with a Zeiss Plan-Apochromat 20x/0.8NA objective lens (ref: 420650-9902-000, Carl Zeiss Microscopy, Germany). Alexa Fluor 647, 546, 488 and DAPI fluorescence were excited with the 639 nm, 561 nm, 488 nm and 405 nm laser lines and collected using the Airy scan 2 detector with the following detection wavelengths: 655-720 nm, 573-627 nm ,499-557 nm and 422-497 nm, respectively. Images were acquired sequentially, in bidirectional mode, with a pixel time of 2.08 µs, at zoom 1.7,using the tile option, in SR mode with a pixel size of 0.062 µm X 0.062 µm, in 16 bit. Airy scan images were processed automatically in 2D, and stitched using Zen Blue 3.1 (Carl Zeiss Microscopy, Germany) software and saved in CZI file format.

Stage 30 embryos were fixed in freshly prepared 4% PFA for 1 hour at RT. Embryos were washed with PBS for 3 × 5 minutes and depigmented using 3% H_2O_2 and 0.8% KOH in DEPC-H_2O for 20 minutes. Next, embryos were incubated in Solution 1 of the Deep-Clear method[34] for 1 hr at 37°C with shaking at 500 rpm on a tabletop shaker. Embryos were washed with 5x SSC-0.1% Tween-20 (SSCT) for 3 × 15 minutes at 500 rpm. Probe hybridization and amplification were done according to the manufacturer's protocol (Molecular instruments, USA). Embryos were mounted in Solution-2 of deep-clear protocol and incubated for 1 hr at room temperature to match the refractive index before imaging. Images were acquired using a Spinning-disk confocal unit (CSU-W1, 25 µm pinhole diameter, Yokogawa, Japan) on a Nikon inverted Ti-Eclipse microscope stand (Nikon Instruments Inc., Japan), equipped with a Nikon CFI Plan Apochromat Lambda D 10X/0.45 for whole embryo imaging and a Nikon CFI Plan Apochromat Lambda 20X/0.75 lens for the cropped images. Alexa 647, mCherry and Dapi fluorescence were excited with a 640 nm, 561 nm and 405 nm laser line and collected using a 405/488/561/640 dichroic beamsplitter (Yokogawa) with a ET705/72 (Chroma), ET605/52 (Chroma) and ET436/20 (Chroma) emission filter respectively. Images were acquired at 2048X2048 pixels in 16 bit with a Scientific CMOS Zyla 4.2 (Andor Technology, United

Kingdom) controlled with NIS AR 5.41 (Nikon Instruments Inc., Japan) software and saved in Nd2 or Tiff file format. A maximum intensity projection (MIP) of the z-stack was created using Fiji software. Images were acquired using Nikon Ti-E Yokogawa CSU-W1 / Nikon C2+ confocal microscope with 10X Nikon CFI Plan Apochromat Lambda (NA-0.45) lens for whole embryo imaging and with 20X Nikon CFI Plan Apochromat Lambda (NA-0.75) lens for the cropped images. A maximum intensity projection (MIP) of the z-stack was created using Fiji-ImageJ software.

Vertebrae transplantation

To visualise vertebrae, 5 cm animals of both *d/d* and *CAGGs:eGFP* animals were soaked in 0.05% alizarin red S (ARS) for 30-60 minutes and rinsed several times in fresh tap water[40]. Subsequently axolotls were anaesthetised in 0.03% benzocaine (Sigma) and prepared for surgery. Using a SZX16 stereo-microscope (Olympus, Tokyo), equipped with a Zyla sCMOS camera (Andor, Belfast) and an X-cite Xylis LED illuminator (Excelitas Technologies, Waltham) fluorescent microscope to visualise vertebrae, and fine forceps 3-4 vertebrae were carefully extirpated from the *d/d* host. Subsequently 3-4 vertebrae were dissected from the donor and grafted into the previously prepared host. Axolotls were left to recover for 2 weeks, after which tails were amputated through the GFP^+ donor vertebrae and lineage traced for a total of 6 weeks.

scRNAseq of mature and blastema tissue

Tissue dissociation, and preparation for scRNAseq was performed as previously described[6]. In brief, tail tissue was dissociated in a dissociation solution (0.1 units/µL DNaseI, 35 µg/mL Liberase TM in 0.8X PBS) using fine forceps to continuously tear the tissue apart for approximately 30-40 minutes. Sample is passed over a 70 µm Filcon filter (BD# 340605)

and spun down at 2000 rcf for 3 minutes. Cells are carefully resuspended in 0.8X PBS and flow sorted with a FACSAria III (BD Biosciences, Franklin Lakes) using an 85 µm nozzle into 200 µL of RLT buffer (Qiagen, RNA mini kit). Single cell sequencing was done according to manufacturer's instructions (10x genomics, Pleasanton) with v3 chemistry being used for all the experiments with the exception of library 65766 where v2 was used. 5' 10x sequencing was used to capture viral barcode labelled samples with minor modifications[41]. 10 pmol of a GFP specific reverse primer and a 10x specific overhang was spiked into the 5' 10x droplet capture mix. An aliquot of the cDNA is amplified using a truseq read1 forward primer and a GFP specific reverse primer with a nextera read2 overhang. This generates amplicons to perform targeted sequencing and can be integrated with the data from the 5' 10x single cells transcriptome analysis.

Viral barcode library production, infection, and clone calling

Viral barcode labelling was performed by modifying existing 3rd generation foamy virus vector systems based on the celltag principle with modifications[13,15,42]. A 9nt random barcode sequence and a defined 6nt tag was cloned shortly behind the TSS of a puc2MD9-SFFV:GFP-WPRE[42] transfer vector plasmid using Gibson assembly. In brief SFFV:GFP-WPRE plasmids were amplified in two separate fragments using primers with barcode containing overhangs, PCR products were DpnI treated and purified using para-magnetic beads, and cloned using Gibson assembly. Immediately after assembly, samples were heat-shock transformed into stellar competent cells (Takara Bio, Kusatsu). SOC outgrowth was performed for 1 hour at 37˚C, and a 5 µL aliquot was serially diluted and plated on ampicillin agar plates to estimate cloning efficiencies reaching a minimum complexity of $2*10^6$ colonies. Foamy virus was produced using Lenti-X 293T (Takara Bio, Kusatsu) generated as previously described and frozen at -80˚C for long term storage[42]. Viral titers were determined by infecting

HT-1080 cells with serial dilutions of previously frozen viral preps in multi-well plates. 4 days later plates were imaged using a Cell Discoverer 7 (Zeiss, Jena) and GFP^+ cells were scored to estimate viral titers. Concentrated virus was injected into either mature or blastema tissue using glass capillary needles and a PV830 picopump (WPI, Sarasota) using standard protocols[42]. We flow-sorted GFP^+ cells at 21 dpa in two independent experiments and performed scRNAseq using the 5' 10X platform. Barcoded GFP transcripts were recovered by adapting existing gRNA capture and direct sequencing approaches[41]. Pooled libraries will be deposited with addgene (puc2MD9-SSFV:V1-GFP-WPRE, puc2MD9-SSFV:V2-GFP-WPRE, puc2MD9-SSFV:V3-GFP-WPRE). Primers used to generate constructs are listed in table 1.

Clonal analysis of viral barcode labelling

Non-paraxial mesodermal lineages were excluded from the analysis. Barcodes were extracted and filtered (present in ≥ 2 cells and ≥ 2 UMIs per cell). Clones were called using a Jaccard similarity index of ≥0.45. The axolotl has a measured cell cycle ranging from 53-103 hours[7,43]. Assuming an average cell cycle time of 100 hours; Tails infected with V1 4 days prior to amputation would have undergone 5.04 rounds of cell division, V2 4 days after amputation could have undergone 4.08 rounds of division, V3 7 days after amputation could have undergone 3.36 rounds of cell division. This corresponds to maximum clone sizes of 32.89 (V1), 16.91 (V2), and 10.27 (V3) respectively. Reducing the Jaccard similarity index threshold results in unusually large clones being called which is inconsistent with the known cell division rate during tail regeneration.

Called clones are displayed in a matrix to visualise clonal diversity (Fig. 2f). Here each row represents a specific cell type, and each column represents a clone. To ensure clones were not promiscuously called we computationally combined the two independent replicates and in this mixed population called clones using a Jaccard similarity index of ≥0.45. 47/48 clones

were exclusively made up of cells originating from either replicate 1 or replicate 2, while only a single hybrid clone called made up of 2 cell from replicate 1 and 1 cell from replicate 2 (Sup. Fig. 2i). This suggests that with a Jaccard similarity index of ≥0.45 we can reliably call clones in these data-sets and achieve a false discovery rate of 3.6%.

scRNAseq of embryonic tissue

Axolotl embryos were staged to stage 25-35[44] and tail buds were dissected in 0.8X PBS under a SZX10 stereo-microscope (Olympus, Tokyo). Individual cells were manually collected using P20 pipettes and prepared for smart-seq2 using standard protocols in a 96-well format[45].

Processing and analysis of scRNAseq data

The quality of the cDNA and resulting sequencing libraries were checked by bioanalyzer (High Sensitivity DNA Kit, Agilent, 5067-4626). The libraries were sequenced using Illumina NovaSeq PE150. Sequencing reads were mapped against the axolotl genome[46,47] using STARsolo (STAR 2.7.5a)[48] with parameters mimicking Cell Ranger's (10x Genomics) transcript counting strategy. Fluorescent marker genes were manually added to the references. Seurat v3.1[49] was mainly used for data analysis. Ribosomal protein genes and pseudogenes were excluded for downstream analyses. Generally, cells with a very high or a low number of transcript counts, as well as those with a high mitochondrial transcript proportion were excluded (see exact information in Table 2). When clusters were identified to be still driven only by low RNA counts in a first screening, cells belonging to these clusters were removed. To enrich connective tissue cells, some contaminating cell types such as epidermal cells were identified and removed after an initial screening using Seurat. Seurat's built-in scale function was used to regress out differences in nUMI and percent.mt. Scaling and PCA were always performed on all genes. Uniform Manifold Approximation and Projection (UMAP) was

applied to the top principal components (PCs) for visualisation. Data integration was performed for the blastema and the uninjured data sets, respectively, using Seurat's built in Harmony[50] function to reduce the confounding effect of different batches and samples. Running Harmony was performed with default parameters on the library level and 100 input PCs for blastema integration and 50 input PCs for uninjured integration. The integration of the uninjured and blastema data sets was performed using Harmony on the library level with slightly adapted parameters (PC=50, theta=2, lambda=9).

Trajectory analysis

To infer directionality of developmental trajectories during regeneration we used scVelo[51]. To infer the ratio of spliced and unspliced reads using scVelo, the sequencing data was mapped with STARsolo and its built-in 'Velocyto' mode[48] to obtain spliced and unspliced count matrices, respectively. Afterwards scVelo was used following the standard procedure (https://scvelo.readthedocs.io/en/stable/VelocityBasics) without dynamical mapping. 50 input PCs were used for the moments calculation to obtain the arrow velocity graph. Velocity diffusion time was calculated and used as a prior for the identification of links and directionality between the clusters by PAGA.[52] The diffusion pseudotime presented in the figure was calculated using the standard diffusion map algorithm[53] in order to set the asomitic cells manually to 0 and demarcate them as the starting population.

Statistical analysis

Statistical analysis was performed using a combination of PRISM (version 9.0.0), and R where appropriate. Student's t-test, ANOVA etc. was performed using PRISM. MZ Structural break test was performed in R.

Acknowledgments

We would like to acknowledge the valuable services provided by the animal facility, Molecular Biology Services, and BioOptics facilities at the IMP; Histology, and NGS sequencing core facilities at the Vienna Biocenter; animal core, light microscopy facility (LMF; RRID:SCR_019166) and Comparative Genomics and Data Science Core at MDIBL which is supported by the Maine INBRE grant (GM103423) from the NIH-NIGMS.

Competing interests

The authors declare no competing interests exist.

Funding sources

WM is supported by an FWF Lise Meitner fellowship DOI 10.55776/M2444, and project grant DOI 10.55776/P34841; TD is supported by a scholarship from the Ecole Normale Supérieure de Lyon, France; EMT is supported by an ERC AdG 742046 of the European Commission; PM is supported by grants from NIH (NIGMS-COBRE-P20GM104318), and DFG- 527098031.

This research was funded in whole, or in part, by the Austrian Science Fund (FWF) DOI 10.55776/M2444 and 10.55776/P34841 and the European Research Council (ERC) RegenMems 742046. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Author contributions

Conceptualization: WM, TG, FF, PM, EMT

Methodology: WM, TG, FF, MP, YST, TYL, TK, JW, JFF, PM,

Investigation: WM, TG, FF, SCP, MP, VSJ, JFF, TD, YST, TYL, TK

Visualisation: WM, TG, FF, PM, BT

Funding acquisition: WM, TD, EMT, PM

Project administration: WM, EMT, PM

Supervision: EMT, PM, BT, DA

Writing - original draft: WM, TG, EMT, PM